Machine Learning in Production Systems Design Using Genetic Algorithms

Abu Qudeiri Jaber, Yamamoto Hidehiko and Rizauddin Ramli

Abstract—To create a solution for a specific problem in machine learning, the solution is constructed from the data or by use a search method. Genetic algorithms are a model of machine learning that can be used to find nearest optimal solution. While the great advantage of genetic algorithms is the fact that they find a solution through evolution, this is also the biggest disadvantage. Evolution is inductive; in nature life does not evolve towards a good solution but it evolves away from bad circumstances. This can cause a species to evolve into an evolutionary dead end. In order to reduce the effect of this disadvantage we propose a new learning tool (criteria) which can be included into the genetic algorithms generations to compare the previous population and the current population and then decide whether is effective to continue with the previous population or the current population, the proposed learning tool is called as Keeping Efficient Population (KEP). We applied a GA based on KEP to the production line layout problem, as a result KEP keep the evaluation direction increases and stops any deviation in the evaluation.

Keywords—Genetic algorithms, Layout problem, Machine learning, Production system.

I. INTRODUCTION

To create a solution to a specific problem in machine learning, the solution is constructed from the data or by use a search method to find nearest optimal solution. Genetic algorithm (GA) is stochastic search which is often used in machine learning application. An effective GA representation and meaningful fitness evaluation is the keys of the success in GA applications. The effectiveness of GA comes from their simplicity as robust search algorithm as well as from their ability to discover good solutions rapidly for difficult high-dimensional problems. GA is useful and efficient in the following cases: ① When the search space is large, complex or poorly understood. ② When the domain knowledge is scarce or expert knowledge is difficult to encode to narrow the search space. ③ If no mathematical analysis is available. ④ When the traditional search methods fail. The advantage of the GA approach is that it can handle arbitrary kinds of constraints and objectives; all such things can be handled as weighted components of the fitness function, making it easy to adapt the GA scheduler to the particular requirements of a very wide range of possible overall objectives. GA is one of the wide applicability of knowledge-learn problem-solving tools that can provide an approach [1], [2]. While the great advantage of GA is the fact that they find a solution through evolution, this is also the biggest disadvantage. Evolution is inductive; in nature life does not evolve towards a good solution but it evolves away from bad circumstances. This can cause a species to evolve into an evolutionary dead end. In order to reduce the effect of this disadvantage a learning tool is add to the GA to decide whether is effective to continue with the generated population or the current population, this learning tool is called hereafter as keeping efficient population (KEP).

II. GENETIC ALGORITHMS

The GA paradigm has been proposed to solve a wide range of problems [2][3][4]. GA has been successfully applied to optimization problems in diverse fields and it is differs from other search techniques which depend on natural genetic evaluation process. GA starts with an initial set of solutions selected randomly called population. A suitable encoding for each solution in the population is used to allow computation of the fitness. The solution set in the population, called as chromosome or individual, represents a solution to the optimization problem. Each individual contains a number of genes. The individuals in the initial population are evaluated to measure its fitness. To create the next population, new individuals are formed by either merging two individuals from the current population using a crossover operator or modifying an individual solution using mutation operator. Based on the individuals’ fitness, the individuals to be included in the next population are then probabilistically selected from the set of individuals in current population. The iteration, called a generation is continued until the fitness reaches its maximum value, with the hope that strong parent will create a fitter generation of the children. The best overall solution becomes the candidate solution to the problem. To create the next generation GA based on three operations: Selection, crossover and mutation.

A. Basic Genetic Algorithm

The outline of the basic GA is described below.

1. [Start] Generate random population of n individuals (suitable solutions for the problem).
2. **[Fitness]** Evaluate the fitness $f(x)$ of each individual $x$ in the population.

3. **[New population]** Create a new population by repeating following steps until the new population is complete.
   - **[Selection]** Select two parent individuals from a population according to their fitness (the better fitness, the bigger chance to be selected)
   - **[Crossover]** With a crossover probability cross over the parents to form a new offspring (children). If no crossover was performed, offspring is an exact copy of parents.
   - **[Mutation]** With a mutation probability mutate new offspring at each locus (position in individual).
   - **[Accepting]** Place new offspring in a new population

4. **[Replace]** Use new generated population for a further run of algorithm

1. The size of the population of solutions, which will determine the diversity of solutions that the algorithm evaluates, and therefore the chance of an optimal solution being found [5].
2. The ability of the algorithm to maintain diversity in the population [6]. The flow chart of GA operations is shown in figure 1.

**B. Using Genetic Algorithms as the Search Procedure**

Genetic algorithms typically maintain a constant-sized population of individuals which represent samples of the space to be searched. Each individual is evaluated on the basis of its overall fitness with respect to the given application domain. New individuals (samples of the search space) are produced by selecting high performing individuals to produce ”offspring” which retain many of the features of their ”parents”. The result is an evolving population that has improved fitness with respect to the given goal. The solution process will generate a large amount of information that can be used to carry out some form of learning.

**C. Keeping efficient population**

In the conventional GA, the set of possible solutions from one population are taken and are used to construct a new population. This is motivated by a hope that the new population will be fitter than the old one. Evolution evolves away from
bad circumstances. This can cause a species to evolve into an evolutionary dead end. This research aims to keep the evaluation direction increase and stop any deviation in the evaluation. In order to achieve this, we proposed a KEP. KEP is a new learning tool which can be included into the genetic algorithms operations to keep the efficient population. KEP is added to the GA to decide whether is effective to continue with the previous population or the current population. The main difference between the conventional GA and GA with KEP is that the conventional GA are used the current population for learning to produce the next generation. In contract, GA with KEP is first compare the current population with the previous population, and then decides whether is effective to continue with the previous population or the current population. By this way KEP is stopped any deviation in the evaluation.

III. PRODUCTION LINE LAYOUT PROBLEM

One of the problems encountered of the design and implementation of a flexible transfer line (FTL) is the layout of the FTL in the restricted area. The layout of the FTL has an important impact on production cost. The efficient layout design of the FTL can reduce the cost of material handling by at least 10-30%. Depending on the production system, between 15% - 70% of the total production cost can be attributed to material handling [7]. Thus, the FTL layout is important to reduce production cost. The FTL layout problem is identified as a NP-complete problem [8], and many heuristic approaches have been developed to solve this problem for near optimum solution [7] – [19]. However, despite the many existing methodologies regarding the FTL layout problem, almost all approaches studying this problem neglect important operational details such as the buffer size between the machines and restricted areas in the plant area. In this paper operational details including the movement of the products between machines are considered, in addition to the restricted areas in the site. In order to find FTL layout including machines and a material handling path between each pair of machines, an efficient FTL layout design procedure called a One by One Layout Method (OOLM) in conjunction with GA is proposed. The OOLM generates an efficient solution for a set of irregularly shaped machines through a restricted plant area. The OOLM is not limited to a single static environment, but is highly flexible, within the plant structure. A CAD system is linked to the proposed OOLM to draw the FTL layout.

The problem can be described mathematically as follows:

Given:
- A set of K irregular-shaped machines and their dimensions;
- Spaces and machine allocation limitations;
- The plant restrictions, such as plant columns, walls, and any other restrictions in the plant area;
- The buffer size vector \( [B_1, B_2, \ldots, B_K] \).

Determine: FTL layout design. So as to minimize \( A_F \). With the following condition
\[ A_F \leq A_P \text{ and } D_k \text{ is located at or is close to the FTL end point.} \]

Where
- \( A_F \): The required area needed to create the new FTL layout.
- \( A_P \): The available plant area.
- \( D_k \): The last machine drop point.

IV. SEARCH A FTL LAYOUT USING GA

In order to find an efficient layout of the FTL components through the available plant area, a GA is included into the OOLM. The OOLM is conducted using the following algorithm. Before describing the algorithm, the notations and terms used in algorithm are defined.

A. Notations

\( K \): Number of machines in the FTL.
\( B_{i+1} \): Buffer allocation between machines \( i \) and \( i+1 \).
\( S_i^c \): Side length of each cell.
\( L_i \): Overall machine length.
\( W_i \): Overall machine width.
\( V \): Overall plant length.
\( l \): Side length of the buffer space.
\( Blocked-cells(i) \): Number of cells included in blocked area \( i \).
\( Machine-cells(i) \): Number of cells included in machine layout \( i \).
\( Blocked-areas \): Number of blocked areas throughout the extended area.
\( C_{start} \): Cell through the plant area that represents the FTL start point.
\( C_{end} \): Cell through the plant area that represents the FTL end point.
\( c_{i-j}^{l} \): Location of cell \( j \) in a machine \( i \) cell map relative to the machine pick point.
\( c_{l}^{b} \): Location of cell \( k \) in blocked area \( i \).
\( c_{l}^{d} \): Cell location of the machine \( i \) drop point through the plant area.

B. Definitions

In order to find an efficient layout of the FTL components through the available plant area, a GA is included into the OOLM. The OOLM is conducted using the following algorithm. Before describing the algorithm, the notations and terms used in algorithm are defined.

[Definition 1] Machine pick point, \( P_i \): The machine pick point is the point where the parts input into machine \( i \).
[Definition 2] Machine drop point, \( D_i \): The machine drop point is the point where the parts leave machine \( i \).
[Definition 3] Machine cells group, \( \mathcal{M}_{cells}^{i} \): The machine cells group is the set of cells that has been mapped to machine \( i \) relative to the \( P_i \) cell.
[Definition 4] Buffer-cell ratio, \( B_{m-c} \): The buffer–cell ratio is...
the smallest integer number greater than or equal to the division of the buffer side length to the cell length; $B_{m-c}$ can be calculated by using equation 2.

$$B_{m-c} = \frac{l}{S_l}$$  \hspace{1cm} (1)

[Definition 5] FTL Components: FTL components include the machines and buffer spaces.

[Definition 6] FTL Component location, $\text{COM}^i_{\text{location}}$: The component location is the location of the next component with respect to the current component in the FTL. The component can be located in one of four locations with respect to the current component (up, down, right or left).

[Definition 7] Extended plant area, $A_{\text{plant-ext}}$: The extended plant area is the area of the rectangle that passes through some of the plant outside wall and includes all plant facilities and sub walls.

[Definition 8] Blocked area: The blocked area is the area in the plant that has been restricted by some obstacles such as plant columns, partitions, and sub areas used by built stations, etc. The blocked areas are not usable for allocation of machines or buffers. In addition to the blocked areas in the plant area, the areas in the extended area that are not included within the plant area are assumed to be blocked areas.

[Definition 9] Block area cells group, $B\text{A}^i_{\text{cells-group}}$: The blocked area cells group is the set of cells that have been mapped to the blocked area $i$.

[Definition 10] Machine cells map, $M^i_{\text{map}}$: The machine cells map is the set of cells that have been mapped to the machine layout. The location of each cell in the machine map is defined relative to the machine pick point.

[Definition 11] Machine drop point possible locations, $C_{\text{drop}}^i(j)$: The machine drop point possible locations are the locations of the machine drop points with respect to a given machine pick point. As follows [19], the machine can be located at 0, 90, 180 or 270, thus $1 \leq j \leq 4$. To describe the machine drop point’s possible locations, the following example is introduced.

Example: Based on machine $i$ with its pick and drop points as shown in Figure 2.

The possible locations of the drop point are cells 1, 2, 3 and 4 in Figure 3.

![Fig. 3 Possible locations of machine drop point](image)

[Definition 12] Buffer space possible locations, $B\text{S}^i_{\text{drop}}(j)$: The buffer space $i$ possible locations are the locations of the buffer space $i$ with respect to the location of the component $i-1$. To describe the buffer space possible locations, the following example is introduced.

Example: Consider machine $i$ in Figure 2 and assume that the buffer size next to this machine is 5. The possible locations of all buffers are shown in Figure 4 below. The possible locations of the first buffer space are the cells coded by the number 1, and the possible locations of the second buffer space are the cells that are coded by the number 2, and so on.

![Fig. 4 The possible locations of the buffers](image)

[Definition 13] FTL area, $A_F$: The FTL area is the area used by the FTL components.

[Definition 14] Plant area, $A_P$: The plant area is the available factory area.

C. OOLM Algorithm

Step 1: Define the matrix that represents the extended plant layout area as follows:

Following [18], the plant in this study is defined as uniform squares connected together and denoted in this paper as cells as shown in Figure 2. The restricted areas (obstacles) including the plant columns and any other obstacles are indicated on the plant layout. The start point and the end point of the FTL locations are given and are represented on the plant layout. The plant model can be defined by using the following steps:

1. Draw the smallest rectangle that passes through the outside walls to include all plant facilities and sub walls. The area covered by this rectangle is denoted in this paper as the extended plant area.
2. Divide the extended plant area into uniform cells that are equal to each other and connected to each other.
3. Define a matrix that represents the extended plant layout area as shown in equation 1,
where \( C(i,j) \) represents the cell located in row \( i \) and column \( j \).

Step 2: Set \( C(k) = C(i,j) \) \( \forall k = 1,2,\ldots,(mn) \), \( i = 1,2,\ldots,m \) and \( j = 1,2,\ldots,n \).

Step 3: Read the FTL start point.
Set \( C_{start} = C(i,j) \), \( i \in [1,\ldots,m] \), \( j \in [1,\ldots,n] \)
and \( C(i,j) \notin B_{A_{cells-group}} \) \( \forall k = 1 \ldots \) blocked - areas.

Step 4: Read the FTL end point.
Set \( C_{end} = C(i,j) \), \( i \in [1,\ldots,m] \), \( j \in [1,\ldots,n] \)
and \( C(i,j) \notin B_{A_{cells-group}} \) \( \forall k = 1 \ldots \) blocked - areas.

Step 5: Define the cells map for all machines, \( M_{i}^{cells-group} \) relative to the \( P_{i} \) cell.
Set \( M_{i}^{cells-group} = \{ C_{i}^{\text{start}}, C_{i}^{\text{end}}, \ldots, C_{i}^{\text{machine-cells}}(i) \} \) \( \forall i = 1,\ldots,K \).

Step 6: Define the drop cell for machine \( i \), relative to the \( P_{i} \) cell.
Step 7: Define all blocked areas cells group through the extended plant area, \( B_{A_{cells-group}} \) by \( A_{plan-ext} \).
Set \( B_{A_{cells-group}} = B_{A_{cells-group}} \) \( \forall i = 1,\ldots,K \).

Step 8: Set the first machine pick point at the FTL start point.
Set \( P_{1} = C_{start} \).

Step 9: Find all possible locations of the first machine drop points, \( cd_{i}^{1} \), as described in definition 11.

Step 10: Randomly select one of the possible locations of the machine drop points.

Step 11: Find all possible locations of the first buffer space with respect to the machine drop point as described in definition 12.

Step 12: Randomly select one of the possible locations of the first buffer space.

Step 13: Find all possible locations of the next buffer space with respect to the current buffer space location and randomly select one of these locations.

Step 14: Repeat step 13 to find one location for each space of buffer size.

Step 15: Find all possible locations of the next machine pick points, \( cp_{i}^{1} \), with respect to the last buffer space and randomly select one of these locations.

Step 16: Repeat steps 13 - 15 to find one of the possible locations of all machines and spaces of all buffers.

Step 17: Apply the GA operations to find the best locations for all of the FTL components.

Step 18: Calculate the FTL area by using equation 3.

\[
A_T = \sum_{i=1}^{K} (L_i \times W_i) + \sum_{i=1}^{K} (B_i \times l_i^2)
\]

(3)

Step 19: Carry out the following rule.
If the evaluate equation 4 is satisfied,
\[
A_P - A_T \geq 0
\]
Then: draw the FTL, Else: read new buffer size vector \([B_1, B_2,\ldots, B_K]\).

V. GENETIC ALGORITHMS OPERATIONS FOR PRODUCTION LINE LAYOUT PROBLEM

The GA operations for the production line layout problem are described in sections V.A ~ V.D.

A. Encoding

One of the important jobs to use GA is how to express a chromosome. One of the main difficulties in encoding this problem is that each pair of components in the FTL is located relative to each other, which it means that each component’s location should be specifically identified in the individual. This research adopts each component’s location in the FTL as the gene, each machine in the FTL represented by two genes one for the machine pick point location and the other for the machine drop point location. The other component’s locations are represented by one gene for each. The conventional GA operations are generally based on an individual with a similar gene size. In the case that we are studying here, namely FTL layout, it is difficult to use a similar gene size in individuals. This is because the number of possible component locations is not equal for all components. The number of possible location for next component is decided according to the state of locations up, down, left and right of the current component. For example if component \( i \) is located at cell \((I,J)\), then the number of possible locations, \( NPL_i \), of component \( i+1 \) is as follows:

\[
NPL_i = \begin{cases} 
0 & \text{if the constraints 1,2,3, and 4 are satisfied} \\
1 & \text{if three constraints from constraints 1, 2, 3, and 4 are satisfied} \\
2 & \text{if two constraints from constraints 1, 2, 3, and 4 are satisfied} \\
3 & \text{if one constraint from constraints 1, 2, 3, and 4 is satisfied} \\
4 & \text{if the constraints 1, 2, 3, and 4 are not satisfied} 
\end{cases}
\]

(5)

Constrain 1:
\((I,J+1) \in \{ ch_{i,1}^{1}, ch_{i,2}^{1}, \ldots, ch_{i,\text{Blocked} - \text{cells}(i)}^{1} \} \) \( \forall i = 1,2,\ldots, \text{Blocked - area} \)

Constrain 2:
\((I,J-1) \in \{ ch_{i,1}^{1}, ch_{i,2}^{1}, \ldots, ch_{i,\text{Blocked} - \text{cells}(i)}^{1} \} \) \( \forall i = 1,2,\ldots, \text{Blocked - area} \)

Constrain 3:
\((I+1,J) \in \{ ch_{i,1}^{1}, ch_{i,2}^{1}, \ldots, ch_{i,\text{Blocked} - \text{cells}(i)}^{1} \} \) \( \forall i = 1,2,\ldots, \text{Blocked - area} \)

Constrain 4:
\((I-1,J) \in \{ ch_{i,1}^{1}, ch_{i,2}^{1}, \ldots, ch_{i,\text{Blocked} - \text{cells}(i)}^{1} \} \) \( \forall i = 1,2,\ldots, \text{Blocked - area} \)

In this research, we propose a new individual encode method to express each individual. The new encode method is called a one to one encoding method (OOWEM). The following paragraph describes how to use OOWEM to encode the components in the FTL. The first element in the individual represents the location of the drop point of the first machine in the FTL. The location of the first machine drop point can be
determined with respect to the location of the first machine pick point as described in definition 11. Elements 2, 3, ..., \( G_{B_1+1} \) in the individual represent the location of the spaces of the first buffer. The location of the first buffer space is determined with respect to the location of machine drop point and the location of next buffer space is determined with respect to the current buffer space location and so on. Figure 5 shows the encoding using the OOEM.

The number of genes in each individual, \( T \), is calculated using Eq. (6).

\[
T = \left( 2K - 1 + \sum_{i=1}^{K-1} B_i \right)
\]  

The number of items in the individuals is not limited, which means that any production line with any number of machines and any buffer size between each pair of machines can be dealt with.

\[\text{Step 1: Select cell } U \text{ (the first cell location in the individual) as the first cell location of child } I.\]

\[\text{Step 2: Find the edges after gene } i \text{ in both parents with respect to the first cell location of child } I: \text{ (first cell location, } U_{\text{read}} \text{) read as the cell up to the first cell location and (first cell location, } L_{\text{read}} \text{) read as the cell left to the first cell location, find the cell state of these two edges. If the cell located left to the first cell location is empty (neither restricted nor used by another components) then it will be chosen as the second gene for the child } I, \text{ otherwise, the cell up to the first cell location will be chosen as the second gene for the child } I. \text{ Assuming that the cell up to the first cell location is empty, we select } U \text{ as the next gene of child } I.\]

\[\text{Step 3: Find the edges after gene 2: (second cell location, } L_{\text{read}} \text{), and (second cell location, } L_{\text{read}} \text{). The edges after gene 2 is the same, we select } L \text{ as the next gene of child } I.\]

\[\text{Step 4: Find the edges after gene 3: (third cell location, } D_{\text{read}} \text{), and (third cell location, } U_{\text{read}} \text{). Assuming that the cell up to the third cell location is not empty, we select } D \text{ as the next gene of child } I.\]

\[\text{Step 5: Find the edges after gene 4: (fourth cell location, } R_{\text{read}} \text{), and (fourth cell location, } L_{\text{read}} \text{). Assuming that the cells right and left to the fourth cell location are not empty, we select } D \text{ as the next gene of child } I. \text{ Note that if cell down to gene 4 is also not empty in this case the previous gene (gene 3 in this example) is change to third possible choice which is } L \text{, and continue from this point again.}\]

\[\text{Step 6: Repeat the same rule in previous step to generate all genes of child 1.}\]

We can use the same procedure to generate child 2 as shown in Fig. 6. After crossover, both offspring encode legal path.
D. Mutation

The mutation of our GA system is different from the traditional mutation operator because the gene expression adopts OOEM. Instead of using the traditional mutation operator, we randomly select two mutation points (two genes in one individual) and swap their values. After swap the two genes, the cells locations of the two mutation points and of the genes between the two mutation points are test, if any of these cells belong to block area cells group, the two mutation points are reselected. The new mutation points are selected between the two old mutation points. This process is repeated until the cells belong to the genes between the two mutation points is not restricted. Thus, we still have legal path after swap mutation.

The mutation is carried out using the following steps.

Step 1: Select one individual randomly from the current population.
Step 2: Select two mutation points, MP1 and MP2, randomly.
Step 3: Swap the two mutation genes values.
Step 4: Check the cells belong to the two mutation points and all cells belong to the genes between the two mutation points, if the cells of two mutation point and the cells between them after swap are empty (neither restricted nor used by another components) then go to Step 7, otherwise go to Step 5.
Step 5: Select new two mutation points, but in this time between the current mutation points
Step 6: Go to Step 3.
Step 7: Accept the mutation.

The follow chart in Figure 7 shows the mutations algorithm.

VI. Numerical Experiments

We applied the developed OOEM to some of FTL examples as follows:

A. Applied FTL example: 10 machines and 9 buffers

The FTL we adopted in this example has 10 machines and the 9 buffers. Each buffer space is assumed to be equal of the plant cells, $B_{m-c} = 1$. The machines shapes and specifications are as shown in Figure 8. Table 1 gives the buffer size between each pair of machines in the FTL.
TABLE I
UNITS FOR MAGNETIC PROPERTIES

<table>
<thead>
<tr>
<th>B1</th>
<th>B2</th>
<th>B3</th>
<th>B4</th>
<th>B5</th>
<th>B6</th>
<th>B7</th>
<th>B8</th>
<th>B9</th>
</tr>
</thead>
<tbody>
<tr>
<td>3</td>
<td>5</td>
<td>5</td>
<td>8</td>
<td>10</td>
<td>9</td>
<td>4</td>
<td>6</td>
<td>9</td>
</tr>
</tbody>
</table>

![Fig. 8 Machines shapes and specifications](image1)

**Fig. 8 Machines shapes and specifications**

**B. Results**

The 3D drawing of the FTL layout resulted by the OOLM is drawn by the CAD system as shown in Figure 9.

![Fig. 9 The 3D drawing of FTL layout](image2)

**Fig. 9 The 3D drawing of FTL layout**

**VII. CONCLUSION**

This research introduces a new learning tool which can be included into the genetic algorithms operations to keep the evaluation direction increase and stop any deviation in the evaluation. GA with KEP is first compare the current population with the previous population, and then decides whether is effective to continue with the previous population or determine. The combination of the OOLM and GA will applied to obtain the layout of any FTL in any plant restricted area.

**REFERENCES**


